PREPARED FOR SUBMISSION TO JCAP

How randomly is the Lyα forest sampling the Universe?

Andreu Font-Ribera,[a,b] **Uroš Seljak**[a,b,c,d]

[a]Institute of Theoretical Physics, University of Zurich, 8057 Zurich, Switzerland
[b]Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720, USA
[c]Department of Physics, Department of Astronomy, University of California, Berkeley, California 94720, USA
[d]Institute for the Early Universe, Ewha Womans University, Seoul 120-750, S. Korea

E-mail: font@pyshik.uzh.ch, useljak@berkeley.edu

Abstract. We usually claim that the Lyα forest traces randomly distributed regions of the universe. With the increase of Lyα data and, and the fact that recent studies have started to use pixels even closer to the quasar (up to few h^{-1}Mpc, Slosar et al. 2013), this statement should be revisited and properly quantified. Here we present an analytical study of this effect, using a simple method based on perturbation theory.

Keywords: Lyα forest , BAO, large scale structure

ArXiv ePrint: 13XX.XXXX

Contents

1 Introduction

The Lyα forest community has always claimed that the forest is tracing random positions of the universe, based in the fact that traditional studies restricted their analysis to fairly large separations from the background quasars. For instance, the upper limit in restframe wavelength often used $\lambda_r = 1185\mathring{A}$ corresponds to $\sim 75\ h^{-1}$Mpc at the redshifts of interest, and the lower limit of $\lambda_r = 1041\mathring{A}$ corresponds to $\sim 425\ h^{-1}$Mpc.

This approximation was enough for previous Lyα studies, but with the recent increase of available data from BOSS this statement should be revisited. Moreover, recent studies of the Lyα forest in BOSS have used less conservative data cuts and have extended the definition of the forest up to restframe wavelengths of $\lambda_r = 1210\mathring{A}$, only $\sim 15\ h^{-1}$Mpc in front of the background quasar.

In this study we quantify this effect by developing a simple analytical model based on perturbation theory. In section 2 we present the model and quantify the effect. In section 3 we show the measured correlation function measured at different separations from the background quasars, discuss the results in 4 and conclude in 5.

[AF: *It would also be possible to test the model in simulations. For instance, Martin White has a set of simulations that might be useful here:* $L = 250\ h^{-1}$ *Mpc box,* $N = 2048^3$ *particles (I think...), with haloes already identified and lines of sight generated only on halo positions. I played with them a couple of years ago, so I have already the code to use them. I could talk to him if we think they might be useful. I would not trust the Ly*α *statistics there, since the resolution is not good enough (and its not hydro), but it could be good enough to qualitatively test the model by looking at the correlation measured at different separations from the quasar.*]

2 Effect of the background quasar

As noted above, we can only sample the Universe with Lyα absorption whenever we have a background quasar. Therefore, in a real survey what we are measuring is the conditional probability for $\delta_F(\mathbf{x})$, $\delta_F(\mathbf{r}+\mathbf{x})$, given that there are two quasars at a certain line of sight separation l_1,l_2 from the pixels. This effect will introduce second order corrections to the measured correlation function, since it involves 3- and 4-point functions. This correction are usually very small, but some of the terms become important once we start using pixels that are very close to its background quasar.

In this study we will only consider the presence of one of the quasars, the one closer to any of the pixels in the pair. The effect of the more distant quasar will be obviously smaller. What is the correlation function that we measure if one of the pixels is always at a separation l from the quasar?

Let's split the Universe in tiny cells, such that each cell contain a number of quasars Q equal to either one or zero. Therefore, the mean number of quasars in a cell $\bar{Q}$ will be very small $\bar{Q} \ll 1$. We can now define the fluctuation $\delta_Q = Q/\bar{Q} - 1$, and compute the 3-point function:

$$\langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x})\ \delta_Q(\mathbf{x}+\mathbf{l}) \rangle = \left\langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x})\ \left(\frac{Q(\mathbf{x}+\mathbf{l})}{\bar{Q}} - 1 \right) \right\rangle \tag{2.1}$$

$$= \frac{1}{\bar{Q}} \langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x})\ Q(\mathbf{x}+\mathbf{l}) \rangle - \langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x}) \rangle \tag{2.2}$$

$$= \langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x}) \rangle_Q - \langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x}) \rangle\ , \tag{2.3}$$

where $\langle X(\mathbf{x}) \rangle_Q$ means average over pixels that are at a line of sight separation l from a quasar. In the last step we have used the fact that Q is equal to one with probability $\bar{Q}$ and zero otherwise.

It is easy to see that $\langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x}) \rangle_Q$ is actually what we are measuring, while $\langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x}) \rangle$ is what we wanted to measure instead. Therefore, the effect of the background quasar is equal to the 3-point function $\langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x})\ \delta_Q(\mathbf{x}+\mathbf{l}) \rangle$.

2.1 3-point function

At first order in δ, the 3-point function $\langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x})\ \delta_Q(\mathbf{x}+\mathbf{l}) \rangle$ vanishes. Therefore, we need to go to a higher order in perturbation theory.

We will describe our Lyα fluctuation as

$$\delta_F = b_F^{(1)} \delta_m + b_F^{(2)} \delta_m^2\ , \tag{2.4}$$

where $b_F^{(n)}$ is the n-th order bias parameter. Equivalently we can write the quasar density fluctuation as

$$\delta_Q = b_Q^{(1)} \delta_m + b_Q^{(2)} \delta_m^2\ . \tag{2.5}$$

Assuming that δ_m is Gaussian (so that we can use Wick's theorem), and going to 4-th order in δ_m, we have

$$\langle \delta_F(\mathbf{x})\ \delta_F(\mathbf{r}+\mathbf{x})\ \delta_Q(\mathbf{x}+\mathbf{l}) \rangle = b_F^{(1)} b_F^{(1)} b_Q^{(2)}\ \xi_m(\mathbf{l})\ \xi_m(\mathbf{r}{+}\mathbf{l}) + b_F^{(2)} b_F^{(1)} b_Q^{(1)}\ \xi_m(\mathbf{l})\ \xi_m(\mathbf{r}) = b_F^{(1)} b_F^{(2)} b_Q^{(1)}\ \xi_m(\mathbf{r})\ \xi_m(\mathbf{r}{+}\mathbf{l})\ . \tag{2.6}$$

If we are interested in the effect on the BAO scale, $r >> l$, the last term will be clearly smaller since $\xi(r)$ is a decreasing function of separation. On the other hand, in this scenario $\|\mathbf{r}+\mathbf{l}\| \sim r$ and therefore both terms might have similar contributions.

Note that this picture is very simplistic, since it does not take into account the effect of redshift space distorsions, that might have a similar contribution if $\beta_F \sim 1$.

2.2 Quantifying the effect

The minimum separation to the background quasar depends on what is the maximum restframe wavelength λ_r used in the study. Both are related by:

$$l = \int_{z(\lambda_r, z_q)}^{z_q} dz' \frac{c}{H(z')}, \tag{2.7}$$

where z_q is the quasar redshift, and

$$z(\lambda_r, z_q) = (1 + z_q) \frac{\lambda_r}{\lambda_\alpha} - 1\ . \tag{2.8}$$

Slosar et al. 2013 used $\lambda_r = 1210$ Å, that for a typical quasar at $z_q = 2.4$ implies $l \sim 15\ h^{-1}$Mpc. [AF: *Give some numbers to biases and quantify effect.*]

3 Data

We will now take a look at the correlation function in the BOSS survey, and try to see this effect by looking at the correlation function measured at different separations from the background quasar.

We use DR10 data, with PCA continua (from KG) plus *Mean Transmission Correction* (MTC, basically require $<\delta>= 0$ in each spectrum), so expect distorted correlations. Correlation measured in 46 sub-chunks and error bars computed from scatter.

The correlation is measured in 29 bins in r, of width $dr = 5\ h^{-1}$Mpc, ranging from 5 to 150 h^{-1}Mpc. There are also 3 bins in μ, width $d\mu = 0.333$. Finally, there are 4 bins in line of sight separation from the closest quasar l: $10 < l < 30$, $30 < l < 50$, $50 < l < 70$ and $l > 70$ (in separations h^{-1}Mpc).

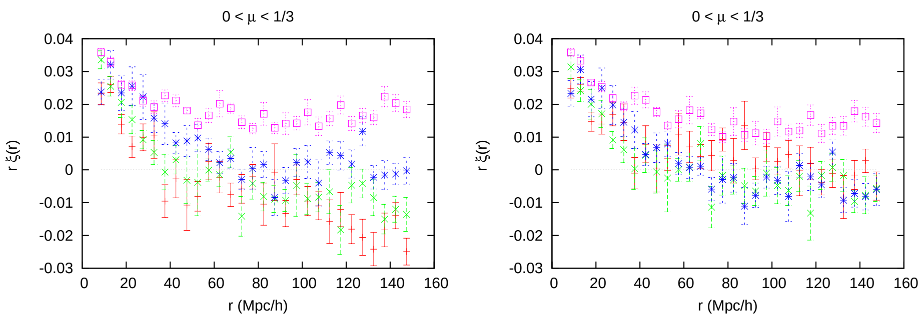

Figure 1. Measured correlation function, in the lowest bins in μ $(0 < \mu < 1/3)$). Each color identifies one of the bins in l, from small to large: red, green, blue, pink. Left panel doesn't correct for the stack in restframe wavelength (errors in the continuum templates).

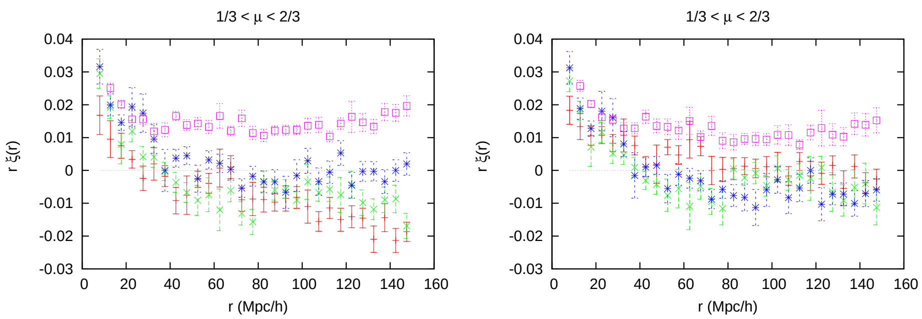

Figure 2. Measured correlation function, in the mid bins in μ $(1/3 < \mu < 2/3)$). Each color identifies one of the bins in l, from small to large: red, green, blue, pink. Left panel doesn't correct for the stack in restframe wavelength (errors in the continuum templates).

Measure correlations are shown in figures 1,2, 3.

[AF: *There is something funny going on in the measured correlations. It looks like the bins with lower restframe wavelengths doesn't go to 0, what doesn't make much sense to me. I'm looking at it right now.*]

4 Discussion

Do we see any effect in the data?

Can the model describe the data? If not, what could be improved?

Can this explain the larger bump in Slosar et al. 2013?

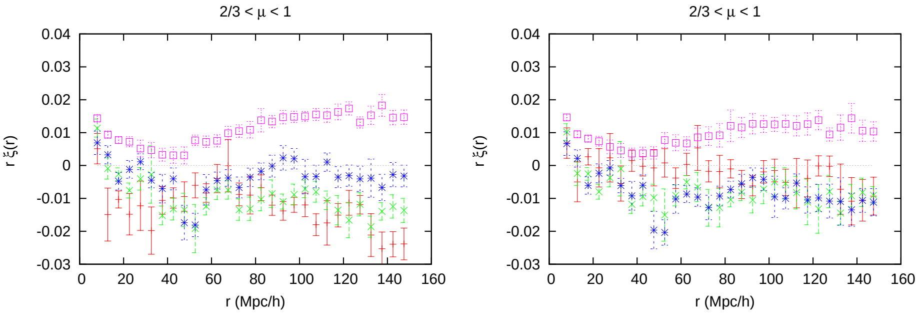

Figure 3. Measured correlation function, in the highest bins in μ ($2/3 < \mu < 1$)). Each color identifies one of the bins in l, from small to large: red, green, blue, pink. Left panel doesn't correct for the stack in restframe wavelength (errors in the continuum templates).

5 Conclusions

summarize and conclude.

[AF: *Appendices are just random notes, nothing interesting in there.*]